MEDICAL RESEARCH LABORATORY REPORT NO. 74

MEDICAL RESEARCH DEPARTMENT

U. S. Submarine Base
New London

THE EFFECT OF INCREASING THE DIFFICULTY OF THE RPA TEST OF NIGHT VISION.

3875-a

Interval Report No. 3
BuMed Project X-268 (Av-156-c)
Field Test of Radium Plaque Adaptometer

Report prepared
by
Lieut.(jg) W. S. Verplanck, H(S), USNR

21 May 1945

APPROVED: C. W. Shilling, (MC) USN, MO-in-C.

THE EFFECT OF INCREASING THE DIFFICULTY OF THE RPA TEST OF NIGHT VISION

Interval Report No. 3
BuMed Project X-268 (Av-156-c)
Field Test of Radium Plaque Adaptometer

Report prepared
by
Lieut.(jg) W. S. Verplanck, H(S), USNR

21 May 1945

APPROVED: C. W. Shilling, (MC) USN, MO-in-C.

THE EFFECT OF INCREASING THE DIFFICULTY OF THE RPA TEST OF NIGHT VISION

SUMMARY

Results obtained on a group of 446 men given both the standard Radium Plaque Adaptometer test and also the same test at a brightness .20 log units lower than the standard, indicate that a much larger percentage of men will necessarily fail the test if the brightness is lowered. Classification based on a more difficult test has the same statistical properties as that based on the regular test, and consequently will doubtless be no more reliable.

Among the several suggestions which have been advanced for the improvement of the Radium Plaque Adaptometer test of night vision, one of the most common has been the addition of one or more lower brightness levels to the present single-brightness-level test. An experiment has been performed to test the effect of such a modification.

Experimental Procedure:

Two Radium Plaque Adaptometers were fitted with easily manipulated extra "neutral" filters of scotopic density .20 plus minus .02 log units. Each man of a group of 219 (Group A) was tested first on the standard plaque by the regular procedure. Immediately thereafter, the test was repeated, with the extra filter in place. The procedure was not altered in any other manner, and the subjects were not informed that they were receiving any but the ordinary test.

Another group of 227 men (Group B) was administered the more difficult test first, with the standard test immediately following. This group served as a check on the possible affect of practice upon the RPA scores, made in the present experiment, and as a source of information on "encouragement" or "discouragement" arising from taking a test of greater or lesser difficulty first.

All tests were performed by three Qualified RPA Operators.

Results:

Results are reported in terms of Level I (regular RPA testing level, approximately 3.9 log uul) and Level II (regular RPA with extra .20 filter added, approximately 3.7 log uul).

Cumulative Frequencies:

Tables IA and IB present the distribution of scores made by Group A on both testing levels, together with cumulative frequencies and cumulative percentages. Corresponding data for Group B are given in Tables IIA and IIB. Tables IIIA and IIIB present percentages failing in each group, and corresponding Chi-Squares on each of the testing levels.

Figure 1 is a graphical presentation of cumulative frequency data. These data in an identical plot on Probability Paper are presented in Figure 1b, which show by the rectilinearity of the curve fitted to the data that the distribution of RPA scores follows a normal distribution. The parallelisim of the two curves indicates that the variability is the same on both testing levels. Deviations at the upper ends of the curves from the fitted straight-lines are the result of the small number of cases represented by each point.

These data indicate that Groups A and B are adequate samples of the general population; there is no significant difference between their performance on either level. The differences in the percentages of failures are within the error of sampling, and no systematic differences in performance which may be attributed to the variation in the order of administration are evident.

Comparison of Performance on Levels II and III.

Scatter plots of scores made on Level I against scores made on Level II are given in Tables IV and V for Groups A and B respectively, together with approximate tetrachoric r's (cosine formula)* about the pass-fail cut. These results are summarized for the combined groups in Table VI A and B.

The performance of both groups on the test at Level II differs significantly from that on Level I. Level II is unequivocally a more difficult test.

* More accurate determinations of this value were not possible, owing to the small number of cases in one cell of the double dichotomy which necessitated a correction.

It should be noted that of the 69 men of both groups who failed the regular RPA test (Level I), only one passes the test at Level II. The conventional cut-off fails an inferior group, without equivocation. Tetrachoric $\overline{r}$'s (cosine formula) compare not unfavorably with others reported on tests of night vision, and conform with statistical expectancy.

Selection of Superior and Inferior Groups:

Table VIIA, B, and C presents the cumulative frequency and percentage of scores made on Level II for each of the following sub-groups of the whole group of 446 men: Men making 10/10 on Level I, men making from 16/20 to 19/20 inclusive on Level I, and men failing on Level I. Figure 2a presents these data graphically, and permits the estimation of the median and standard deviation of scores made by each sub-group. These are included in Table VIIIA. The chi-square test of the hypothesis that each of these sub-groups is a representative sample of the whole group is given in Table VIIIB. Figure 2b is a replot of Figure 2a on Probability Paper. Again we are dealing with normally distributed scores, and with the same variability.

Values of chi-square indicate that performance at Level I predicts performance on Level II with statistical significance. The median scores of the sub-groups differ by almost one standard deviation, a difference which is significant. This finding suggests that those making 10/10 on the RPA test may be classed "superior". This is very probably a procedure of dubious merit, however, since the validity of the RPA is undetermined.

Discussion:

In previous reports*, the statistical basis of the RPA test of night vision has been presented. A test of this type must be based upon the statistical properties of small samples. A sample of 20 trials provides an estimate of the individual's probability of seeing at a fixed brightness level. The limits of accuracy are such that all individuals whose "true" probability of seeing is 0% will score 11/20 and below at the .001 level of confidence. In order to fail, those whose probability of seeing is 30% or less at the same level of confidence, a failing score of 15/20 is required. These statements hold irrespective of the brightness level at which the test may be run. Any Pass-Fail cutting score, then, if it is to be meaningful, must be based not on the percentage of men it is desired to classify as FAIL, but on the statistical probability of a given test score. The percentage of men failed must be manipulated by variation of the brightness level at which the test is administered.

In the present experiment, in which the difficulty of the test was increased by lowering the brightness of the test level by .20 l.u., the percentage of men failing, if the 30% frequency of seeing is desired as a cutting performance, is radically increased from 15.5% to 59.9%. If it is desired only to eliminate those whose frequency of seeing is 0%, the percentage of failures must be increased from 4.9 to 28.8% in order to insure that no such man can pass.

The present testing brightness was selected before these considerations were apparent. It has proven satisfactory, however, and has been used since the test was first designed.

* Report of Five Hundred Subjects Tested and Retested for Fifty Trials Each on the Navy Radium Plaque Adaptometer, Research Division, Bureau of Medicine and Surgery, Washington, D. C., of 8 March 1944.

The same statistical properties of the test sharply limit the correlation which may be found between performances on two levels of difficulty of the same test, and the test-retest correlation as well. The present data confirm the first expectation. No direct data on test-retest reliability are presented, but it may legitmately be inferred that it will behave as expected.

In view of these considerations, it may be asserted that the reliability of the RPA test at a lower brightness level will be the same as that at the present level. Nor will the validity of the test be altered.

Conclusions:

1. The effect of setting up a more difficult test by fitting the RPA with an additional filter of scotopic density .20 plus minus .02 log units, is to produce a test which a larger number of men will necessarily fail. This result is of doubtful desirability in a test of undetermined validity.

2. The regular RPA test, as well as the more difficult test, gives scores failing a normal distribution.

3. The variability of the population is the same in both tests.

4. It is possible to identify, on the basis of the regular RPA scores, three groups of different performances in the more difficult test. Then scores of these sub-groups, who may be called "superior", average" and "inferior" are normally distributed, and the variability of each group averages the same magnitude. Practical application of this finding would be unwise, since the RPA is a test of undetermined validity.

5. No significant purpose will be served by altering the difficulty of the RPA test of night vision. To make the test more difficult would have the result that a larger percentage of the population fail. This is of doubtful desirability in a test of undetermined validity.

TABLE IA

Group A

Distribution of Scores on Level I (Standard RFA TEST)

Score	F	CF	C%
10/10	132	132	60
19/20	14	146	67
18/20	13	159	73
17/20	19	178	81
16/20	13	191	87
15/20	7	198	90
14/20	8	206	94
13/20	3	209	95
12/20	4	213	97
11/20	1	214	98
10/20	0	214	98
9/20	0	214	98
8/20	0	214	98
7/20	1	215	98
6/20	2	217	99
5/20	1	218	99
4/20	1	219	100

TABLE IB

Distribution of Scores on Level II (Extra .20 1.u Filter)

Score	F	CF	C%
10/10	41	41	19
19/20	7	48	22
18/20	8	56	25
17/20	15	71	32
16/20	18	89	41
15/20	18	107	49
14/20	10	117	53
13/20	15	132	60
12/20	21	153	70
11/20	9	162	74
10/20	15	177	81
9/20	12	189	86
8/20	10	199	91
7/20	8	207	94
6/20	7	214	98
5/20	2	216	99
4/20	3	219	100

F -- Frequency
CF -- Cumulative Frequency
C% -- Cumulative Percentage

TABLE IIA

TABLE IIB

roup B

Distribution of Scores on Level I (Standard RFA Test)

Score	F	CF	C%
10/10	123	123	45
19/20	16	139	61
18/20	15	154	68
17/20	19	173	76
16/20	13	186	82
15/20	11	197	87
14/20	5	202	89
13/20	3	205	90
12/20	6	211	95
11/20	4	215	95
10/20	8	223	99
9/20	2	225	99
8/20	1	226	99
7/20	1	227	100

Distribution of Scores on Level II (Extra .20 l.u Filter)

Score	F	CF	C%
10/10	40	40	18
19/20	4	44	20
18/20	10	54	24
17/20	17	71	31
16/20	19	90	39
15/20	27	117	50
14/20	20	137	59
13/20	21	158	68
12/20	11	169	73
11/20	20	189	82
10/20	12	201	87
9/20	8	209	90
8/20	6	215	93
7/20	1	216	93
6/20	4	220	95
5/20	3	223	96
4/20	1	224	96
3/20	2	226	98
2/20	1	227	100

F -- Frequency
CF -- Cumulative Frequency
C% -- Cumulative Percentage

TABLE III

Chi Square Test of Homogenity of
Men in Group A and Group B.

A
Level I

	Group A	Group B	
		Obtained	Predicted*
PASS	191	186	198
FAIL	28	41	29

Chi-Square = 1.45

The Chi-Square test shows that the chances that the performance of Group B on Level I differs from that of Group A are greater than 20, and less than 30 in 100. The groups, then, constitutetwo samples of the same homogenous population.

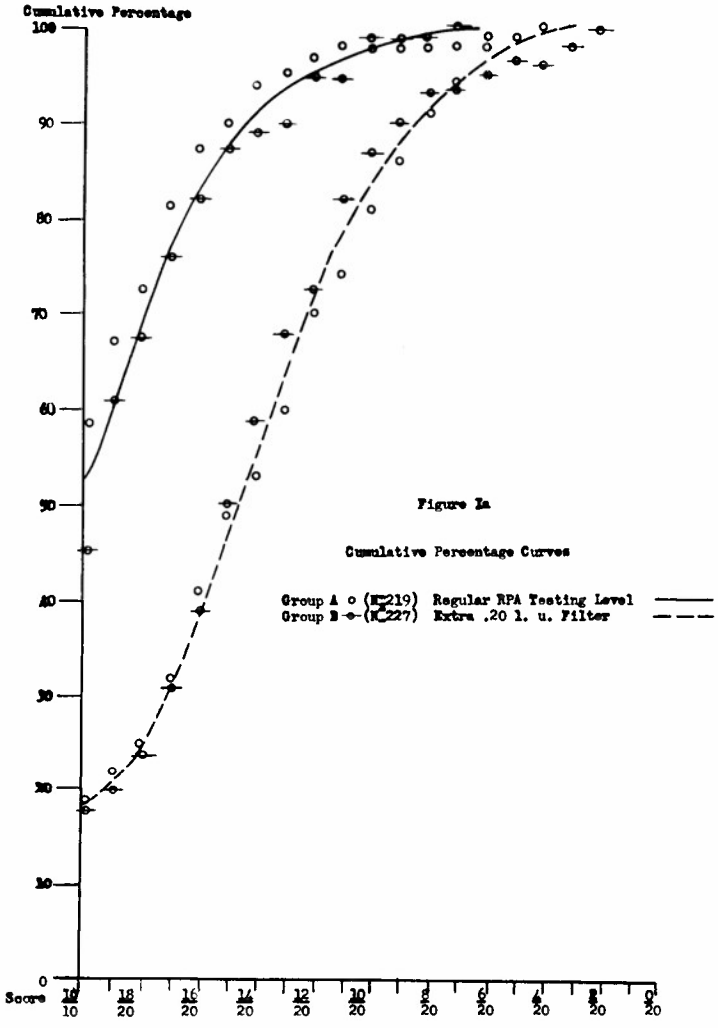
B
Level II

	Group A	Group B	
		Obtained	Predicted*
PASS	89	90	92
FAIL	130	137	135

Chi-Square = 0.10

The Chi-Square test shows that the chances that the performance of Group B on Level II differs from that of Group A are greater than 50, and less than 95 in 100. The groups, then, constitute two samples of the same homogenous population.

* From results of Group A


Cumulative Percentage
100
90
80
70
60
50
40
30
20
10
0
Score
10/10
18/20
16/20
14/20
12/20
10/20
8/20
6/20
4/20
2/20
0/20
Figure 1a
Cumulative Percentage Curves
Group A o (K-219) Regular RPA Testing Level
Group B o (K-227) Extra .20 l. u. Filter

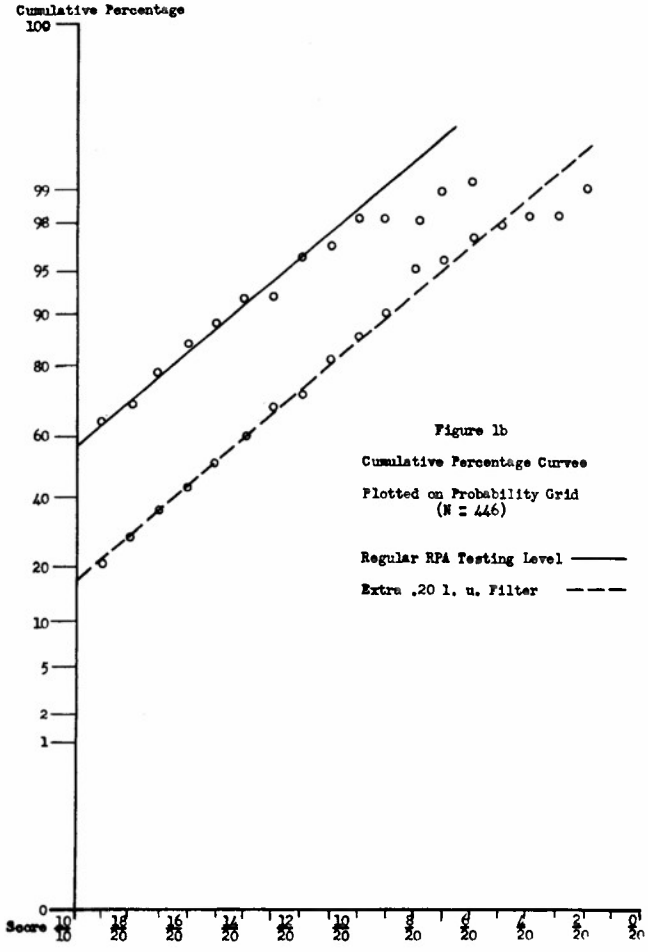


Figure 1b

Cumulative Percentage Curves

Plotted on Probability Grid

(N = 446)

TABLE IV

Scatter Plot -- Group A

Tetrachoric r* about 10/10 -- .56
Tetrachoric r about 16/20 -- .57

Level I (3.9 log u.u.l.) Standard RPA Test

Level II - (3.7 log u.u.l.)

Score	1/20	2/20	3/20	4/20	5/20	6/20	7/20	8/20	9/20	10/20	11/20	12/20	13/20	14/20	15/20	16/20	17/20	18/20	19/20	10/10
10/10																	1	1	4	35
19/20																			1	6
18/20																		1		7
17/20													1				1		1	12
16/20																1	1	1	1	14
15/20															1		2		2	13
14/20														1		1		1		7
13/20															1	3	1	3		7
12/20														1		3	2	1	2	12
11/20														1			2		1	5
10/20												2		1	2	2	2	3		3
9/20													1	1	1	2	3			4
8/20													1	1		1	1	1	1	4
7/20												1		2	1		2	1		1
6/20						1					1	1			1		1		1	1
5/20				1	1															
4/20						1	1													1
3/20																				
2/20																				
1/20																				

* All tetrachoric r's (cosine formula) are approximate, a correction being necessary for the small number of cases in one quadrant.

TABLE V

Scatter Plot - Group B

Tetrachoric r about 10/10 -- .66
Tetrachoric r about 16/20 -- .67

Level I (3.9 log u.u.l.) Standard RPA Test

Level II - (3.7 log u.u.l.) Score	1/20	2/20	3/20	4/20	5/20	6/20	7/20	8/20	9/20	10/20	11/20	12/20	13/20	14/20	15/20	16/20	17/20	18/20	19/20	20/20
10/10																		1	1	38
19/20																		1		3
18/20																			1	9
17/20																1	1	4	1	10
16/20																1	2	1	2	13
15/20															2	2	3	2	6	12
14/20											1				1	1	2	1		14
13/20												1		3	2	1	2	1	2	9
12/20										1	1	1				1	3			4
11/20										2		1	2	1	1		3	2	2	6
10/20										1					1	4	2	1	1	2
9/20									1		1	2			1			1		2
8/20												1	1	1	1	1				1
7/20																	1			
6/20							1			2					1					
5/20								1	1		1									
4/20																1				
3/20										1					1					
2/20										1										
1/20																				

TABLE VI A

PASS-FAIL Double Dichotomy
Level I vs. Level II
(Groups A and B Grouped)

	PASS (Level I)	FAIL (Level I)	TOTAL
PASS (Level II)	178	1	179 (40.1%)
FAIL (Level II) Plus filter	199	68	267 (59.9%)
TOTAL	377	69	446
	(84.5%)	(15.5%)	

Tetrachoric r = .76 (approx.)
(cosine formula)

TABLE VI B

Chi-Square Test of Homogeneity of Data
obtained on Level I and Level II

	Level I	Level II	
	Obtained	Obtained	Predicted
PASS	377	179	377
FAIL	69	267	69

Chi-Square = 208

Chi-Square test shows that the chances are infinitesimal (less than one in more than 100,000) that there is no difference in test performance on Level I and Level II.

TABLE VII

Cumulative Frequencies and Percentages of Scores on Level II on Men Scoring in Each of Three Groups on Level I

Level II	A Group Scoring 10/10 on Level I			B Group Scoring 19/20–16/20 on Level I			C Group Scoring 15/20 or less on Level I		
	F	CF	C%	F	CF	C%	F	CF	C%
10/10	73	73	29	8	8	7	0	0	0
19/20	9	82	32	2	10	8	0	0	0
18/20	16	98	38	2	12	10	0	0	0
17/20	22	120	47	9	21	17	1	1	1
16/20	27	147	58	10	31	25	0	1	1
15/20	25	172	67	17	48	39	3	4	6
14/20	21	193	76	6	54	44	3	7	10
13/20	16	209	82	13	67	55	7	14	20
12/20	16	225	88	12	79	64	4	12	26
11/20	11	236	93	10	89	72	8	26	38
10/20	5	241	95	15	104	85	7	33	48
9/20	6	247	97	6	110	89	8	41	60
8/20	5	252	99	5	115	94	6	47	68
7/20	1	253	99	4	119	98	4	51	74
6/20	1	254	99	2	121	99	8	59	86
5/20	0	254	99	0	121	99	5	64	93
4/20	1	255	100	1	122	100	2	66	96
3/20							2	68	99
2/20							1	69	100

TABLE VIIIA

Medians and Estimated Standard Deviations
of Scores Made by Sub-Groups

	10/10 Group	19/20-16/20 Group	15/20 and less Group
Median Score	17/20	13.5/20	10/20
Standard Deviation	4/20	4/20	4/20

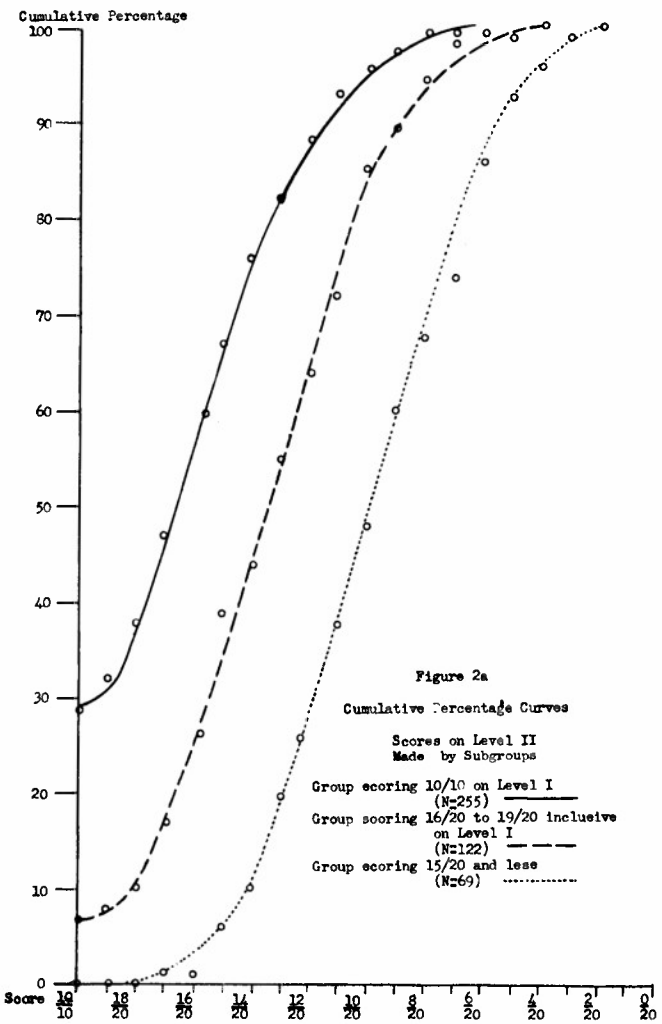
TABLE VIIIB

Chi-Squares of Pass-Fail Cut for Each Sub-Group

	10/10 Group N = 255	19/20-16/20 Group N = 122	15/20 and less Group N = 69
Predicted Fails*	152.9	73.6	39.3
Obtained Fails	108	92	67
Chi-Square **	39.1	13.7	25.7
p - less than	0.0000...	0.0000...	0.0000...

* On basis of performance of whole group (N = 446).

** Chi-square test shows that the probability that performance on level I has not permitted selection of three sub groups of different ability in performance on level II are less than 1 in more than 6000.


Cumulative Percentage
100
90
80
70
60
50
40
30
20
10
0
Score
10/10
18/20
16/20
14/20
12/20
10/20
8/20
6/20
4/20
2/20
0/20
Figure 2a
Cumulative Percentage Curves
Scores on Level II
Made by Subgroups
Group scoring 10/10 on Level I (N=255)
Group scoring 16/20 to 19/20 inclusive on Level I (N=122)
Group scoring 15/20 and less (N=69)

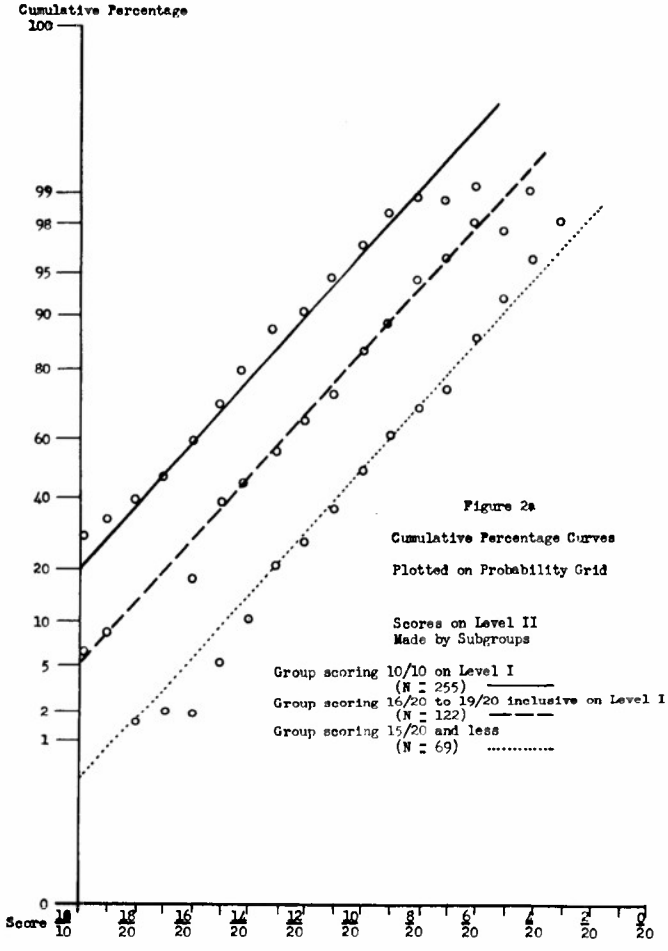


Figure 2a

Cumulative Percentage Curves

Plotted on Probability Grid

Scores on Level II
Made by Subgroups

Group scoring 10/10 on Level I
(N = 255) ————

Group scoring 16/20 to 19/20 inclusive on Level I
(N = 122) — — —

Group scoring 15/20 and less
(N = 69) .....